Bandolier 105

Independent evidence-based health care

Winning the lottery

"There is much luck in the world, but it is luck. We are none of us safe". So said EM Forster nearly 100 years ago. ***Bandolier*** is constantly astonished in its travels by people who appreciate the importance of chance, in, say, winning a lottery or avoiding a car accident, but not in clinical trials. Perhaps it is all down to the way statistics are taught. We should forget probabilities and p-values, and acquaint ourselves with more relevant information, notably how much data do we need to be sure that an observation is not likely to occur just by chance.

Why are people impressed with p-values? The cherished value of 0.05 merely says that a result is not more likely to have occurred by chance than 1 time in 20. Most of us have played Monopoly or other games involved with throwing dice. We will have experienced that throwing two sixes with two dice happens relatively often, yet the chance of that is about 1 time in 36.

Look at it another way. If you were about to cross a bridge, and were told that there was a 1 in 20 chance of it falling down when you were on it, would you take the chance? What about 1 in 100, or 1 in 1000? That p-value of 0.05 also tells you that 1 time in 20 the bridge **will** fall down.

The dice analogy is pertinent, because there are now (at least) two papers that look at random chance and clinical trials, reminding us how often and how much chance can affect results. An older study actually used dice to mimic clinical trials in stroke prevention [1], while a more recent study [2] used computer simulations of cancer therapy.

DICE 1

In this study [1] participants in a practical class on statistics at a stroke course were given dice and asked to roll them a specified number of times to represent the treatment group of a randomised trial. If six was thrown, this was recorded as a death, with any other number a survival. The procedure was repeated for a control group of similar size. Group size ranged from 5 to 100 patients.

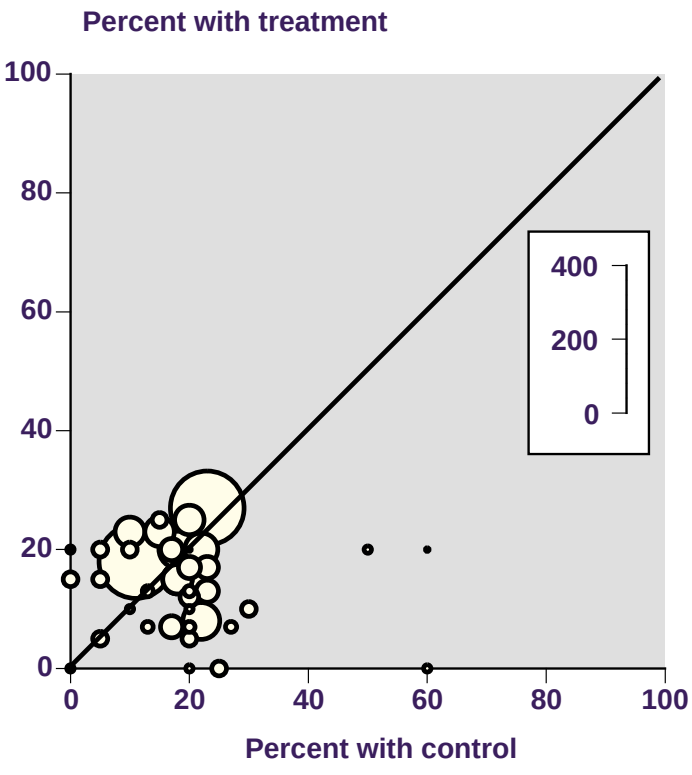
The paper gives the results of all 44 trials for 2,256 "patients". While the paper does many clever things, it is perhaps more instructive to look at the results of the 44 trials. Since each arm of the trial looks for the throwing of one out of six possibilities for standard dice, we might expect that the rate of events was 16.7% (100/6) in each, with an odds ratio or relative risk of 1.

Figure 1 shows a L'Abbé plot of the 44 trials. The expected result is a grouping in the bottom left, on the line of equality at about 17%. Actually, it is a bit more dispersed than that, with some trials far from the line of equality.

The odds ratios for individual trials are shown in Figure 2. Two trials (20 and 40 in total) had odds ratios statistically different from 1. That's one time in every 22 trials, what we expect by chance.

The variability in individual trial arms is shown in Figure 3, where the results are shown for all 88 trial arms. The vertical line shows the overall result (16.7%). Larger samples

Figure 1: L'Abbé plot of DICE 1 trials



In this issue

come close to this, but small samples show values as low as zero, and as high as 60%.

The overall result, pooling data from all 44 trials, showed that events occurred in 16.0% of treatments and 17.6% of controls (overall mean 16.7%). The relative risk was 0.8 (0.5 to 1.1) and the NNT was 63, with a 95% confidence interval than went from one benefit for every 21 treatments to one harm for every 67 treatments (Table 1).

Figure 2: Odds ratios for individual DICE studies, by number in "trial"

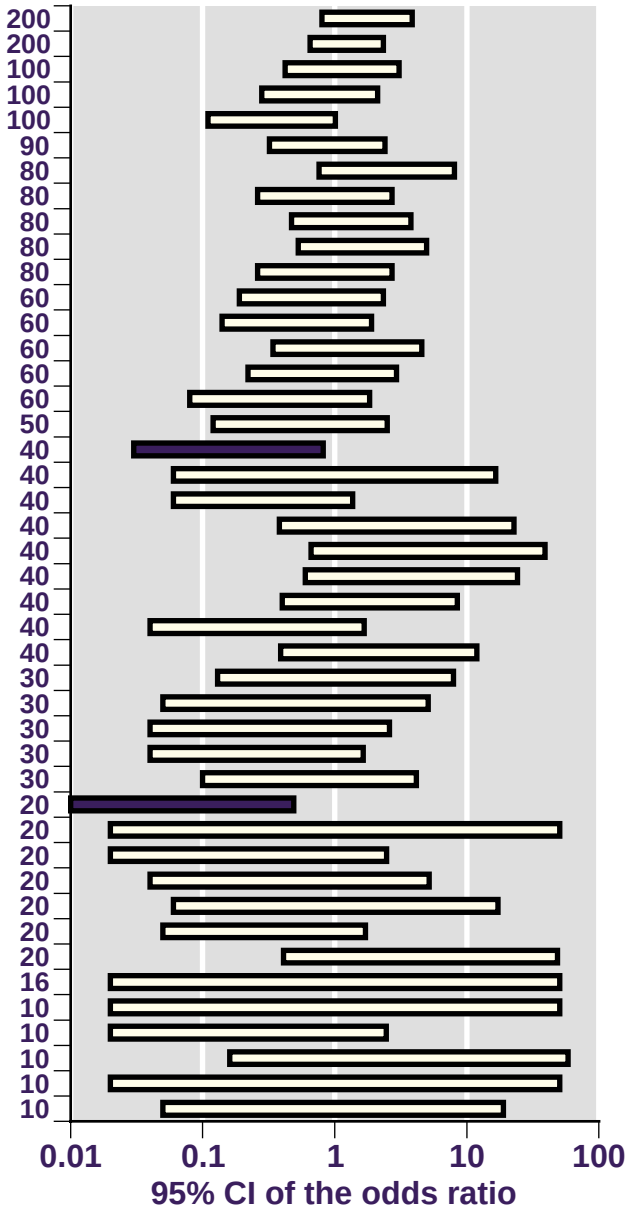
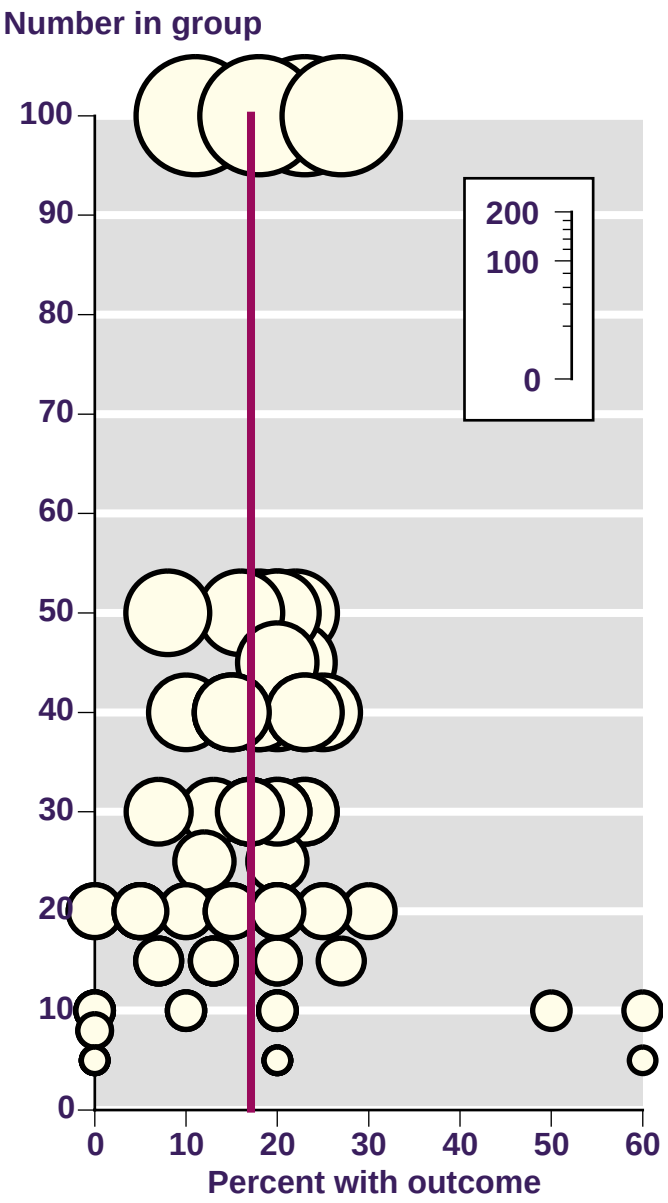


Table 1: Meta-analysis of DICE trials, with sensitivity analysis by size of trial

	Number of		Outcome (%) wih			
	Trials	Patients	Treatment	Control	Relative risk (95% CI)	NNT (95% CI)
All trials	44	2256	16.0	17.6	0.8 (0.5 to 1.1)	62 (21 to -67)
Larger trials (>40 per group)	11	1190	19.5	17.8	1.1 (0.9 to 1.4)	-60 (36 to -16)
Smaller trials (<40 per group)	33	1066	12.0	17.3	0.7 (0.53 to 0.94)	19 (11 to 98)

Figure 3: Percentage of events in each trial arm of DICE "trials"



Many of the experimental DICE trials were quite small, with as few as five per group. The smaller trials, with 40 per group or less, actually came up with a statistically significant result (Table 1). The NNT here was 19 (11 to 98).

DICE 2

Information on the time between randomisation and death in a control group of 580 patients in a colorectal cancer trial was used to simulate 100 theoretical clinical trials. Each time the same 580 patients were randomly allocated to a theoretical treatment or control group, and survival curves calculated [2].

Four of the trials artificially generated had statistically significant results. One was significant at the 0.003 level (1 in 333) and showed a large theoretical decrease in mortality of 40%.

Subgroup analysis was done for this trial by randomly allocating patients to type A or type B, and doing this 100 times. Over half (55%) of the subgroup analyses showed statistical significance between subgroups. The extremes of results were no difference between subgroups, to a result with high significance of 0.00005 (1 in 20,000). In another trial that had bare statistical significance, four of 100 simulated subgroups had statistical significance at the 1 in 100 level.

How much information is enough?

While it is relatively easy to demonstrate that inadequate amounts of information can result in erroneous conclusions, the alternative question, how much information we need to avoid erroneous conclusions, is more difficult to answer. It depends on a number of things. Two important issues are the size of the effect you are looking at (absolute differences between treatment and control), and how sure you want to be.

A worked example using simulations of acute pain trials [3] gives us some idea. Using the same 16% event rate as in DICE 1 as the rate with controls (because it happens to be what is found with placebo), it looked at event rates with treatment of 40%, 50% and 60%, equivalent to NNTs of 4.2, 2.9 and 2.3. The numbers in treatment and placebo group were each simulated from 25 patients per group (trial size 50) to 500 patients per group (trial size 1000). For each condition 10,000 trials were simulated and the percentage where the NNT was within ±0.5 of the true NNT counted.

Table 2: Effect of size and size of effect on confidence of treatment effect

	Percent events with treatment		
	40	50	60
NNT	4.2	2.9	2.3
Group size			
25	26	37	57
50	28	51	73
100	38	61	88
200	55	81	96
300	63	89	99
400	71	93	99
500	74	95	100

With control the event rate was 16

At least

50%	within +/- 0.5
80%	within +/- 0.5
95%	within +/- 0.5

The results are shown in Table 2. With 1000 patients in a trial where the NNT was 2.3, we could be 100% sure that the NNT measured was within ±0.5 of the true NNT; all trials of this size would produce values between 1.8 and 2.8. In a trial of 50 patients where the NNT was 4.2, only one in four trials would produce an NNT within ±0.5; the true value is between 3.7 and 4.7, and three-quarters of trials (or meta-analyses) of this size would produce NNTs below 3.7 or over 4.7.

The study also shows that to be certain of the size of the effect (the NNT, say), we need ten times more information than just to know that there is statistical significance.

Comment

What does all this tell us? It emphasises that the random play of chance is a factor we cannot ignore, and that small trials are more prone to chance effects than larger ones. And it is not just an effect seen in single trials. Even when we pool data from small trials just from rolling dice, as in DICE 1, a meta-analysis can come up with a statistically significant effect when there was none.

High levels of statistical significance can be generated just by the random play of chance. DICE 2 found levels of statistical significance of 1 in 333 for at least one simulated trial, and 1 in 20,000 for a subgroup analysis of that trial.

Not only do we need well-conducted trials of robust design and reporting, we also need large amounts of information if the size of a clinical effect is to be accurately assessed. The rule of thumb is that where the difference between control and treatment is small we need very large amounts. Only when the difference is large (an absolute risk increase or decrease of 50%, affecting every second patient) can we be reasonably happy with information from 500 patients or fewer.

When we see differences between trials, or between responses to placebo, the rush is often to try and explain the difference according to some facet of trial design or patient characteristic. Almost never does anyone ask how likely the difference is to occur just by the random play of chance.

Some things are very unlikely, like winning the lottery. It's the random play of chance, coupled with low downside (£1) and high upside (£millions and our only hope of early retirement to a gin palace) that makes it worthwhile.

References:

1 CE Counsell et al. The miracle of DICE therapy for acute stroke: fact or fictional product of subgroup analysis? BMJ 1994 309: 1677-1681.

2 M Clarke, J Halsey. DICE2: a further investigation of the effects of chance in life, death and subgroup analyses. International Journal of Clinical Practice 2001 55: 240-242.

3 RA Moore et al. Size is everything - large amounts of information are needed to overcome random effects in estimating direction and magnitude of treatment effects. Pain 1998 78: 209-16.

ACUPUNCTURE FOR IDIOPATHIC HEADACHE

Who reads Cochrane reviews in full? ***Bandolier*** has recently conducted a quite unscientific survey of several hundred medical professionals in the UK, and the proportion who have ever read a review in full is under 1%, though perhaps half will have read an abstract. This is a shame. Cochrane reviews are almost always well done, and the biggest disappointment is usually finding how little information there is on some important topics. In highlighting this, Cochrane reviews do a great service.

Cochrane reviews also serve as a terrific resource for teaching. Postgraduate tutors might wish to note this and use them more often. An example is a Cochrane review [1] of acupuncture for idiopathic headache. It concluded:

"Overall, the existing evidence supports the value of acupuncture for the treatment of idiopathic headaches. However, the quality and amount of evidence are not fully convincing. There is an urgent need for well-planned, large-scale studies to assess the effectiveness and cost-effectiveness of acupuncture under real-life conditions."

Note the balance between support on the one hand, and caution on the other. If you read nothing else but this conclusion, would you judge this a worthwhile treatment option? How would you justify it to a priorities forum in your organisation?

Review

A number of relevant databases were searched to identify randomised, or quasi-randomised trials of acupuncture for the treatment of idiopathic headache. Valiant efforts were made to grade trials for quality, validity and appropriateness of the application and method of acupuncture used.

Results

The 26 included trials included 16 trials in patients with migraine (one in children), six in patients with tension-type headache patients, and four in patients with various types of headaches. Eleven trials used standard criteria for headache, and 16 did not. The median treatment period was eight weeks with eight treatment sessions.

Quality and validity scores were generally low. Both randomised and quasi-randomised trials (by alternation or date of birth) were included. Eleven studies (42%) were double blind, and 15 (58%) were open or single blind.

Markers of quality and validity were used to identify better trials:

- a quality score of three or more out of five (9/26 trials),
- an internal validity score of four or more out of six (7/26 trials),
- or a score of 70% or more for the appropriateness of the method and application of acupuncture (10/26 trials).

The overall results of the 26 trials as judged by reviewers are shown in Table 1. A statistically significant positive result in favour of acupuncture was found in only three trials, 12% of the total included.

Seventeen of the 26 included trials used sham acupuncture as control. Three double blind studies were considered to have no interpretable data, either because of baseline differences between treatment groups or high rates of loss to follow up. Table 2 shows the results of trials when information was segregated by different potential biases within the studies. Results were more likely to be negative (no statistical difference between acupuncture and sham acupuncture) in double blind trials, trials of higher reporting quality, trials with higher internal validity, and in larger trials.

Only five of the 26 included trials scored three or more for quality ***and*** scored four or more for validity. One had no interpretable results (Table 3). Only one of these was positive and only one was larger than 50 patients. All had some methodological problems that made them less relevant, like having no clear criteria for patient selection, or being of short duration, or giving no information of use of medicines or intensity, duration or frequency of attacks (Table 3).

Comment

The majority of trials included in this systematic review were small. Often diagnostic criteria or inclusion criteria

Table 1: Overall results for 26 trials of acupuncture for idiopathic headache

Result	Double blind		Not double blind	
	Number	Percent	Number	Percent
No interpretable data	3	12	0	0
Negative: statistically significant	0	0	1	4
Negative: trend	1	4	1	4
No difference	1	4	2	8
Positive: trend	3	12	4	15
Positive: statistically significant	3	12	7	27
Total	11	42	15	58

Positive: acupuncture more effective than control
Negative: control more effective than acupuncture

Table 2: Potential source of bias and confidence in trials of acupuncture for headache

Potential source of bias	Statistical benefit of acupuncture	No benefit of acupuncture
No source of bias considered	7	7
Randomised	6	7
Quasi-randomised	1	0
Double blind trials	3	4
Not double blind trials	4	3
Reporting quality 3 or more	1	4
Reporting quality 2 or less	6	3
Reporting quality 3 or more; validity score 4 or more	1	3
Reporting quality 2 or less; validity score less than 4	5	3
Reporting quality 3 or more; validity score 4 or more; > 50 patients	0	1
Reporting quality 3 or more; validity score 4 or more; < 50 patients	1	3

for patients entering a trial were poorly reported. No information was provided about the severity of headache before administration of study treatment in many of them, and, in some, patient characteristics differed greatly between treatment groups, showing failure of randomisation. Overall, trials were of poor methodological quality, low validity, and often the application of acupuncture was either inappropriate or could not be assessed. The general quality of reporting of information in the trials was poor, especially with regards to description of drop-outs. There are no methodologically rigorous studies assessing clinically relevant outcomes.

Studies of higher quality and validity should be given more weight in systematic reviews because they are more likely to produce reliable results. Just look at how the picture can change when quality, validity and size enter the equation. With no source of bias considered there are seven trials in which acupuncture is statistically better than sham acupuncture. With reporting quality of 3/5 or better, that drops immediately to 1, stays at 1 when validity is added, but drops to zero when we add size. There is **no** evidence that acupuncture is effective. Not a single decent trial.

This review is yet another example of how giving equal weight to poor quality studies can lead to erroneous conclusions. Recommendations cannot be made to implement the use of acupuncture in the treatment of idiopathic headache based on current evidence.

Is there any way that the results of this review could be interpreted differently? Is the reviewers conclusion, that

Table 3: The five "best" trials, of adequate reporting quality and internal validity of acupuncture

Trial	Type of headache	Number of patients	Quality score (max 5)	Validity score (max 6)	Result	Additional comments
1	Migraine	52	3	4	Not significant	Appropriateness of acupuncture 85%; IHS criteria; no information on use of medication, intensity, duration or frequency of attacks.
2	Migraine	30	3	4	Not significant	Appropriateness of acupuncture 45%; Ad hoc definition of headache; did mention information on improved intensity, duration or frequency of attacks. Follow-up data uninterpretable
3	Tension	30	3	5	Not significant	Appropriateness of acupuncture 80%; Ad hoc criteria for headache; no information on intensity or druation of attacks.
4	Tension	25	3	4	Significant benefit of acupuncture	Short duration, poor outcome; limited clinical relevance. Appropriateness of acupuncture 70%; no information on use of medication, intensity, duration or frequency of attacks.
5	Tension	10	5	4.5	No data provided	Baseline group differences in favour of acupuncture; small group size; pilot study. Questionable validity

All trials were randomised and double-blind

"*overall, the existing evidence supports the value of acupuncture for the treatment of idiopathic headaches*" correct? Would you recommend this therapy for your organisation?

High quality, valid trials of adequate size are lacking in this field. There is no reason why studies cannot be adequately randomised or blinded, using International Headache Society recommendations on patient selection and clinically relevant outcomes. Diagnostic criteria should be clearly described, with randomised patients experiencing sufficient pain (moderate or severe) before administration of the intervention. Trials should have a sham acupuncture control. Dosing, duration and follow-up should be adequate and clearly described, as should information on the collection, reporting and severity of adverse effects.

Who should pay? Therapists often bleat about the dearth of research funding for complementary therapy, but then that could be said for many needy areas. The complementary therapy business is huge, costing citizens millions, if not billions of pounds (or euros, or dollars) each year. Therapists (or losers) should pay, as has been done in Holland. Otherwise it would be a bit like a major pharmaceutical company saying they had some interesting chemicals, and could we please pay for their development.

Reference:

1 D Melchart et al. Acupuncture for idiopathic headache (Cochrane Review). In: *The Cochrane Library,* Issue 3, 2002. Oxford: Update Software.

Book reviews

The extraordinary voyage of Pytheas the Greek. Barry Cunliffe, Allen Lane, The Penguin Press, 2001. ISBN 0-713-99509-2. pp 182, £12.99.

This book is not about evidence-based medicine, nor even about medicine, but it is about evidence. It is a remarkable tour-de-force by Barry Cunliffe, the renowned Oxford archaeologist, at the height of his intellectual powers and experience, writing about a topic that clearly excites him.

The Pytheas in question was a native of the Greek colony of Massalia (Marseilles), who made a remarkable journey around the Atlantic fringes 2,300 years ago. He wrote about it in a book, long since lost. He apparently visited Britain (the subtitle of the book is the "man who discovered Britain", though that is a rather classical view), and even travelled as far as Iceland and Denmark.

Exploration outside the Mediterranean was not unknown, and could be quite extensive. The Pharaoh Necho II in about 600 BC arranged for some Phoenician ships to circumnavigate Africa, and several hundred years later Carthaginians penetrated the Atlantic as far as the Sargasso Sea. But the tin isles, and the source of amber, remained something of a mystery.

Tin and amber were motives. The deciding factor was the closure of the straits of Gibraltar by the Carthaginians, and a prolonged state of armed neutrality with western Greek expansion in the Mediterranean.

Cunliffe sets out to recreate Pytheas' book, "*On the Ocean*", from fragments quoted by other writers in later centuries, and vituperous arguments between those who thought Pytheas a great man, and those who considered him a charlatan. What makes this book useful is the way Cunliffe deals with evidence, from textual analysis, agreements or disagreements between texts and archaeology, and even mathematics.

If you love history, this book is a must for its history alone. If you are not a historian, you'll love it for its sense of enquiry and fun. Whatever, you'll learn a little something about using evidence appropriately.

Getting health economics into practice. Edited by David Kernick. Radcliffe Medical Press, 2002. ISBN 1-85775-575-8. pp358, £ not known.

There is lots in this book that is good, but it is curiously unsatisfying. It's not a "curate's egg" problem of being good in parts, because there is nothing bad about any of it. Perhaps the sense of unease comes from what is missing or what is not dealt with in as exciting a way as it could be.

Some examples. The prologue is entitled – better to be vaguely right than precisely wrong. Nice stuff from John Maynard Keynes. But Keynes didn't run a PCT. What technique might I use if I did run a PCT. Operational research is one such. Invented in Oxford by Watson Watt for the better use of radar, it won the war and helped make US industry great. It can be complicated, but is often simple, and helps incorporate economics and efficacy in decision-making. Not mentioned in this book.

Again, programme budgeting and marginal analysis (PBMA) can be really exciting, and especially innovative in healthcare. It might be seen as an ideal technique for healthcare decision-making, and there is talk of ministers becoming interested. Here it is dealt with only from the literature. This isn't very encouraging, but very little good management is published because journals simply are not interested. At least there is an appendix of studies we can get our hands on, some of them very good.

Perhaps that is the problem. Slightly too hands off for the folk to whom it is directed, those involved in planning, commissioning and delivering healthcare. Maybe the fault is in the discipline itself, a bit pointy-headed academic and airy-fairy, when what is needed is a robust sleeves-rolled-up, no-nonsense handbook about how to make things better. A sort of Haynes Manual for the NHS.

Since we won't ever get that, we have to make do with what we can get. This book **will** help. It covers most of the bases, is mostly written in a lively style, and had large doses of reality. Individual chapters, like Chris Newdick's chapter on patients' rights, rationing, and the law are simply terrific. But I'd like to see an accompanying volume on real problems where evidence, economics and management have all been used to make a difference, so that we can all have the same performance as the best.

RURAL GPS: GETTING THEM AND KEEPING THEM

Most of us live in towns or cities. Some of us are lucky enough to live in the country (though that is ***Bandolier***'s subjective view). Recruiting GPs to towns and county is hard enough, especially now in the UK. The general shortage of doctors, early retirement, large older cohorts of GPs coming up to retirement, and the demands for more hospital specialists all mean that recruiting GPs anywhere will be difficult.

In the USA, even though some of those special factors do not apply, recruitment to rural practices has been difficult, and is getting more difficult. A literature review [1] has examined factors that might influence recruitment and retention of rural GPs.

Review

The literature search was extensive, and limited to the years 1990-2000 in order to be contemporary. Examining bibliographies and communication with selected authors extended the search. The definition of rural was that of original authors. The analysis examined factors before medical school, in medical school, and postgraduate factors.

Results

There were 21 studies, all cohorts or cross-sectional studies, often with controls, often with multivariate analysis, often of high quality, and often large (the biggest involving almost 600,000 US physicians). The main results are shown in Table 1, which is no more than an overall summary of some detailed studies.

In the main, having a rural upbringing and having the intention to practice as a GP increased recruitment to rural practices. In medical school specialised programmes giving experience in rural primary care were helpful, and postgraduate rotations with rural experience and obstetric training improved recruitment. Some of these also influenced retention.

Comment

No summary can give full justice to this review. One particular facet was that it described several specialist programmes to increase physician activity to areas underserved by doctors. This would be fertile ground for getting to grips with other areas of difficulty, like inner-city provision of primary care.

Of course there will be other issues beyond choosing the right candidates for medical school, and training them in particular ways in medical school and afterwards. House prices and local education provision are two that spring to mind. But those societal factors apart, this paper would be an interesting read for those organising medical school and postgraduate teaching programmes.

Reference:
1 RG Brooks et al. The roles of nature and nurture in the recruitment and retention of primary care physicians in rural areas: a review of the literature. Academic Medicine 2002 77: 790-798.

Table 1: Main finding of studies examining recruitment and retention of rural GPs, according to pre-medical school, medical school and postgraduate factors

	Number of studies	Recruitment	Retention
Pre-medical school	6	Rural upbringing and intention to practice as a GP increase rural recruitment	No factor clearly linked with retention
Medical school	15	Specialised programmes for experience in rural primary care, and specialised curriculum, plus physician shortage programmes increased likelihood of recruitment	Specialised programmes for experience in rural primary care, and specialised curriculum increased likelihood of retention
Postgraduate	6	Residence programmes with more rural rotations and obstetric training	Rural rotation and residency emphasising underserved health care, and preparedness for small-town living increased retention

ASPIRIN IN LOW RISK INDIVIDUALS

The question of whether prophylactic aspirin protects individuals at low risk of cardiovascular disease keeps being asked (though the exact dose at which it effective keeps being overlooked). ***Bandolier*** 86 carried a review examining the risks and benefits of aspirin use that looked at both coronary events prevented and harmful bleeds produced. The balance tipped from benefit to harm when the annual risk of a cardiovascular event was below 1%. Primary prevention is probably worthwhile at coronary risks of 1.5% a year or more, risks and benefits are balanced at an annual risk of 1%, and aspirin use is unsafe when the risk is 0.5% or less.

A new analysis of studies in low risk patients shows that aspirin use in such individuals has little or no effect on all-cause mortality.

Review

The review sought studies where aspirin was used prophylactically in low risk individuals, and where the outcome reported was all cause mortality. Low risk was defined as having no more than one of a list of risk factors, including hypertension, high cholesterol or LDL, family history, smoking, diabetes, age over 45 years in men and 55 in women, angina, and past cardiovascular events. A widespread search included MEDLINE, Cochrane, previous reviews and the Internet.

Results

No study included only low risk patients. Two randomised trials (US physicians, British doctors) and one cohort study (US nurses) met the inclusion criteria, though grouping low risk individuals with higher risk individuals known to benefit from aspirin use. The aspirin doses were 325 mg every other day, 500 mg daily, and 1-15 aspirin tablets per week, respectively.

In the two randomised trials, the only significant benefit produced was for myocardial infarction in the US physicians study, and that disappeared when combined with the British doctors study (Table 1). The nurses cohort study had significantly increased rates of heart attack, stroke and morality associated with aspirin use.

EDITORS

Andrew Moore Henry McQuay

Pain Relief Unit
The Churchill, Oxford OX3 7LJ

Editorial office: 01865 226132
Editorial fax: 01865 226978
Email: andrew.moore@pru.ox.ac.uk
Internet: www.ebandolier.com
ISSN 1353-9906

Comment

While there is no doubt of the benefits of aspirin use in individuals at high risk of cardiovascular events, there is no evidence that it is beneficial in those at low risk. It can be argued that there is insufficient evidence, rather than evidence of absence of an effect. That argument is justified, but against it could be levied the counter-argument that these studies were biased in favour of aspirin efficacy by including individuals likely to benefit.

The balance of benefit, harm and dose in low risk individuals is another story that waits to be told. Almost half of the British physicians stopped because of gastrointestinal adverse effects on 500 mg a day. But what is the lowest effective dose of aspirin? Would the arguments of benefit and harm be changed by doses of 75 mg daily, or lower? If higher doses have no discernible effect in low risk individuals, would we expect lower doses to be any better?

The practical point is that, for now, there is no evidence that aspirin is of any value in benefiting cardiovascular outcomes or mortality in low risk individuals. There will probably be some harm.

Reference:

1 JM Boltri et al. Aspirin prophylaxis in patients at low risk for cardiovascular disease: a systematic review of all-cause mortality. Journal of Family Practice 2002 51: 700-705.

Table 1: Description and results of individual trials examining prophylactic use of aspirin in individuals at low risk of cardiovascular disease

	US physicians	UK doctors	US nurses
Description			
Study architecture	RCT	RCT	Cohort
Number	11,037	3,429	35,048
Duration (years)	5	6	6
Outcome and odds ratio (95%CI)			
Myocardial infarction	**0.6 (0.5 to 0,7)**	1.0 (0.7 to 1.2)	**2.3 (1.9 to 2.9)**
Stroke	1.3 (0.9 to 1.6)	1.2 (0.8 to 1.7)	**1.8 (1.4 to 2.4)**
Cardiovascular mortality	1.0 (0.7 to 1.3)	0.9 (0.7 to 1.2)	**1.7 (1.2 to 2.3)**
Total mortality	1.0 (0.8 to 1.2)	0.9 (0.7 to 1.1)	**1.8 (1.6 to 2.1)**